

PENGEMBANGAN SOAL PILIHAN GANDA SISTEMIK PADA MATA PELAJARAN KIMIA SMA/MA KELAS XI SEMESTER 2

Erfan Priyambodo dan Marfuatun

Jurusan Pendidikan Kimia FMIPA Universitas Negeri Yogyakarta, erfan@uny.ac.id

Abstrak

Pada setiap proses evaluasi pembelajaran, tidak terkecuali pada pembelajaran kimia, pendidik menggunakan berbagai alat evaluasi yang berupa soal untuk mengukur prestasi belajar peserta didik. Salah satu bentuk soal yang lazim digunakan adalah soal pilihan ganda, dimana mempunyai kelemahan yaitu kurang mampu mengungkap konsep-konsep kimia secara komprehensif. Untuk mengatasi kelemahan tersebut, dikembangkan suatu model soal pilihan ganda sistemik untuk mata pelajaran kimia SMA/MA Kelas XI Semester 2. Soal yang dikembangkan dibatasi pada materi larutan penyangga dan hidrolisis garam sejumlah 30 butir soal. Soal divalidasi oleh 5 ahli materi kemudian dilakukan penilaian kualitas kepada 15 guru kimia SMA/MA. Dari hasil penilaian reviewer, soal pilihan ganda sistemik mempunyai kriteria sangat baik ditinjau dari aspek bahasa, materi dan konstruk. Dari hasil ujicoba kepada 164 siswa SMA/MA, 29 butir soal dinyatakan valid dengan reliabilitas sangat bagus.

Kata Kunci: soal pilihan ganda sistemik

PENDAHULUAN

Kimia merupakan salah satu mata pelajaran di SMA/MA khususnya pada peminatan IPA. Sebagaimana mata pelajaran yang lain, pada mata pelajaran kimia juga dilakukan proses evaluasi. Proses evaluasi merupakan salah satu cara untuk mengukur tingkat keberhasilan proses pembelajaran. Sistem evaluasi yang baik akan mendorong pendidik untuk menentukan strategi yang belajar yang tepat untuk peserta didik, sehingga motivasi peserta didik untuk belajar akan menjadi lebih baik (Djemari Mardapi, 2008 : 5). Evaluasi digunakan untuk membuat keputusan mengenai peringkat dan penempatan peserta didik, kebutuhan-kebutuhan instruksional, dan rencana pembelajaran (Chiheb, dkk, 2011).

Pada setiap proses evaluasi, pendidik menggunakan berbagai alat evaluasi yang berupa soal untuk mengukur prestasi belajar peserta didik. Adapun soal yang digunakan oleh guru antara lain tipe soal pilihan ganda, soal menjodohkan, soal essay, dan lain-lain. Secara umum, soal-soal tersebut dikategorikan menjadi 2 tipe, yaitu *paper based question* dan *computer based question*.

Salah satu bentuk soal untuk mengukur hasil belajar siswa adalah soal pilihan ganda. Soal tipe ini mencakup 2 hal pokok, yaitu soal (*problem*) dan sejumlah alternatif jawaban (Zimmaro, 2004). Soal pilihan ganda dapat digunakan untuk mengukur berbagai tingkatan pengetahuan kognitif (taksonomi Bloom), yaitu dari tingkat pengetahuan sampai ke tingkat evaluasi. Soal pilihan ganda mempunyai beberapa kelebihan dan kekurangan (Zimmaro, 2004). Kelebihan dari soal pilihan ganda antara lain:

1. Dapat mengukur hasil pembelajaran dari tingkatan yang sederhana sampai ke tingkat yang kompleks (pengetahuan sampai dengan evaluasi)
2. Lebih sistemik dibandingkan bentuk soal yang lain, dan memuat tugas-tugas yang jelas
3. Dapat digunakan untuk mengukur suatu sampel dalam jumlah yang besar
4. Alternatif-alternatif jawaban yang salah dapat digunakan sebagai informasi untuk tes diagnostik
5. Skor tidak terlalu terpengaruh oleh "terkaan" peserta didik dibandingkan dengan soal jawaban benar-salah
6. Skor lebih reliabel dibandingkan soal essay
7. Memudah untuk proses pemberian skor
8. Dapat mencakup materi pembelajaran yang cukup banyak

Adapun kelemahan dari soal pilihan ganda adalah:

1. Pada penyusunan yang dilakukan secara baik, membutuhkan waktu yang banyak
2. Seringkali susah untuk menyusun jawaban pengecoh yang masuk akal
3. Tidak efektif untuk mengukur pembelajaran yang menggunakan pendekatan *problem solving*, kemampuan peserta didik untuk mengorganisasi, dan mengemukakan ide-idenya.
4. Skor dapat dipengaruhi oleh kemampuan peserta didik untuk membaca soal
5. Terkadang mengandung lebih dari satu jawaban yang benar.

Soal pilihan ganda yang selama ini digunakan belum mampu mengungkap konsep-konsep kimia secara komprehensif. Umumnya, soal-soal tersebut hanya mengkondisikan peserta didik untuk mengingat kembali pengetahuan yang sesuai dengan topik yang diujikan dan mengaplikasikan pernyataan-pernyataan yang ada (Fahmy dan Lagowski, 2006). Hal ini menyebabkan peserta didik tidak mampu berfikir secara konstruktif, sehingga peserta didik cenderung melupakan konsep-konsep dasar kimia yang telah mereka pelajari sebelumnya. Secara jangka panjang, hal ini tidak baik untuk peserta didik. Jika mereka mudah melupakan konsep-konsep dasar kimia, maka tidak tercapai tujuan pembelajaran kimia.

Soal pilihan ganda sistemik merupakan salah satu pengembangan pendekatan pembelajaran sistemik yang dicetuskan oleh Lagowski (2008). Sistemik yang dimaksud adalah penyusunan konsep-konsep pembelajaran yang saling berhubungan di dalam suatu bagan siklus. Pada soal pilihan ganda sistemik, setiap pilihan jawaban dapat memuat tiga sampai lima konsep-konsep kimia yang saling berhubungan.

Pada soal pilihan ganda sistemik, setiap pilihan jawaban merepresentasikan hubungan tiga atau lebih antara konsep-konsep kimia. Soal pilihan ganda sistemik ini dapat mengevaluasi cakupan materi yang cukup luas, karena mampu menyajikan konsep-konsep kimia secara komprehensif. Oleh karena itu, pengembangan soal pilihan ganda ini sesuai untuk dikembangkan pada mata pelajaran kimia di SMA, karena di SMA cakupan materi kimia sangat banyak dengan topik yang berbeda namun sebenarnya saling berkaitan. Melalui pengembangan alat evaluasi ini, diharapkan peserta didik lebih termotivasi dan terkondisikan untuk berfikir secara konstruktif dan komprehensif.

METODE

Pengembangan soal evaluasi pembelajaran kimia berbasis pendekatan sistemik mengadaptasi prosedur pengembangan Gall, *et al* (2003), yaitu :

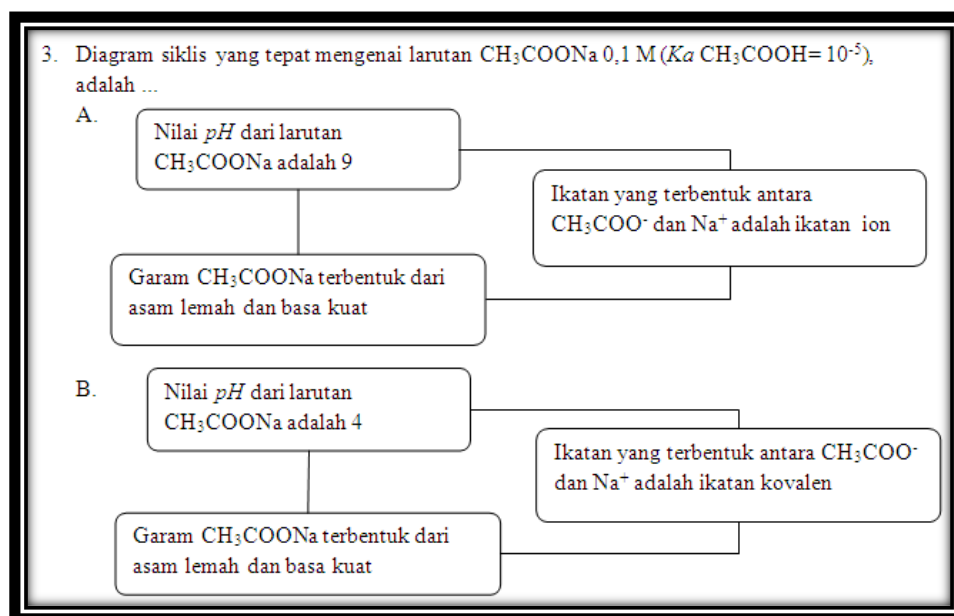
1. Research analysis, needs assessment, and proof of concept
2. Product planning and design
3. Preliminary product development
4. Preliminary field testing
5. Product revision
6. Main field testing
7. The final product revision

Pada penelitian ini, soal pilihan ganda sistemik dibatasi pada materi larutan penyangga dan hidrolisis garam. Kualitas soal pilihan ganda sistemik dilakukan melalui analisis terhadap angket penilaian kualitas soal pilihan ganda sistemik yang telah dinilai oleh responden (15 guru kimia SMA/MA). Analisis data kualitas soal pilihan ganda sistemik dilakukan secara deskriptif. Data juga diperoleh dari skor siswa mengerjakan soal evaluasi berbasis pendekatan sistemik yang di analisis validitas butir soal, dan reliabilitas soal evaluasi menggunakan model aplikasi Rasch (Bambang Sumintono dan Wahyu Widhiarso, 2014).

HASIL DAN PEMBAHASAN

Pada penelitian ini dihasilkan 30 butir soal pilihan ganda sistemik pada materi larutan penyangga dan hidrolisis garam. Soal pilihan ganda berbasis pendekatan sistemik berisikan 5 buah alternatif jawaban, dimana masing-masing alternatif jawaban memuat 3 konsep kimia yang saling berkaitan.

Soal pilihan ganda sistemik maupun instrumen yang telah disusun diujikan kepada 5 dosen ahli materi untuk di-*review*. Hasil dari *review*, digunakan untuk memperbaiki beberapa kesalahan yang masih muncul di soal ataupun instrumen untuk menilai soal. Setelah direvisi, soal dinilai oleh 15 guru kimia SMA/MA untuk mengetahui kualitas soal yang disusun. Adapun contoh soal pilihan ganda sistemik yang dikembangkan dapat dilihat pada Gambar 1 berikut.



Gambar 1. Contoh soal pilihan ganda sistemik yang dikembangkan

Soal yang sudah divalidasi oleh ahli materi kemudian dinilai kepada 15 guru kimia untuk mengetahui kualitasnya. Hasil penilaian tersebut tersaji pada Tabel 1 berikut.

Tabel 1. Kualitas (persentase keidealan) soal pilihan ganda sistemik berdasar penilaian *reviewer*

Aspek	Persentase Keidealan (%)
Bahasa	86,20
Materi	84,94
Konstruk	84,66
Rerata	85,27
Kategori	Sangat Baik

Berikut ulasan mengenai masing – masing aspek yang terdapat di dalam angket yang digunakan sebagai instrumen penilaian soal berbasis pendekatan sistemik.

Aspek bahasa

Aspek bahasa dalam soal evaluasi ini meliputi tiga butir indikator, yaitu kaidah bahasa Indonesia yang baik dan benar, bahasa yang komunikatif, dan bahasa yang tidak menimbulkan persepsi ganda. Secara keseluruhan penilaian kualitas soal evaluasi aspek bahasa sangat baik. Beberapa guru hanya menyarankan untuk memperbaiki beberapa narasi pada pertanyaan dalam soal sehingga mudah dipahami oleh siswa.

Aspek materi

Aspek materi dalam soal evaluasi belajar ini meliputi tujuh butir indikator, yaitu kesesuaian dengan Kompetensi Dasar, kemampuan soal dalam melatih siswa berfikir

konstruktif, muatan konsep kimia yang benar, soal memiliki satu pilihan jawaban yang benar, soal memiliki tingkat kesulitan yang beragam, pilihan jawaban bersifat homogen dan logis, dan pilihan jawaban memuat konsep yang saling berkaitan. Secara keseluruhan penilaian kualitas soal evaluasi aspek materi sangat baik. Namun, pada beberapa butir soal perlu diteliti kembali konsep kimia yang benar sesuai dengan saran dan masukan yang diberikan oleh reviewer.

Aspek Konstruk

Aspek konstruk dalam soal evaluasi belajar ini meliputi sepuluh butir indikator, yaitu pokok soal dirumuskan secara jelas, rumusan pokok soal dan pilihan jawaban merupakan pernyataan yang diperlukan, pokok soal memberi petunjuk ke arah pilihan jawaban yang benar, pokok soal sebagian besar tidak mengandung pernyataan yang bermakna ganda, panjang rumusan pilihan jawaban pada masing-masing pokok soal relatif sama, soal tidak mengandung pernyataan "semua pernyataan jawaban benar" atau "semua pernyataan jawaban salah", gambar, tabel, dan grafik relevan dengan pokok soal, setiap pokok soal yang dibuat tidak bergantung dari pilihan jawaban soal sebelumnya, jenis huruf dapat dibaca dengan jelas, ukuran huruf dapat dibaca dengan jelas. Secara keseluruhan penilaian kualitas soal evaluasi aspek konstruk sangat baik. Namun, pada beberapa butir soal perlu diperbaiki beberapa pilihan jawabannya sesuai dengan saran dan masukan yang diberikan oleh reviewer.

Soal pilihan ganda berbasis pendekatan sistemik merupakan alat evaluasi yang baru dikembangkan. Menurut Wand da Brown dalam Wina Sanjawa (2008 : 335), evaluasi sebagai "...refer to act or process to determining the value of something" atau evaluasi mengacu kepada suatu proses untuk menentukan nilai sesuatu yang dievaluasi. Dalam hal ini proses pembelajaran kimia yang akan dievaluasi dengan model evaluasi yang baru yaitu menggunakan soal yang pilihan jawabannya menggunakan diagram siklik yang memuat beberapa konsep pembelajaran kimia yang telah dipelajari sebelumnya.

Pada penelitian ini, soal yang telah diuji ke siswa dinilai dengan skor dikotomi asli yaitu skor satu (1) untuk jawaban yang benar dan skor nol (0) untuk jawaban yang salah. Jumlah siswa yang terlibat dalam ujicoba ini sejumlah 164 siswa kelas XI. Selanjutnya yang akan dianalisis adalah validitas dan reliabilitas. Data hasil uji coba dianalisis menggunakan aplikasi *Winsteps* dengan model analisis *Rasch*.

Validitas

Validitas suatu instrumen evaluasi adalah derajat yang menunjukkan dimana suatu tes mengukur apa yang hendak diukur (Sukardi, 2011:31). Pada penelitian ini untuk mengetahui validitas butir soal menggunakan aplikasi *Winstep* dengan model analisis *Rasch*. Contoh hasil validitas butir soal untuk kelas XI tersaji pada Tabel 2.

Untuk mengetahui apakah suatu *item* valid, maka harus memenuhi salah satu dari kriteria MNSQ, ZSTD, dan Pt Mean Coor seperti yang sudah disebutkan pada metode penelitian. Kriteria yang berwarna hijau pada Tabel 2 menandakan bahwa *item* tersebut memenuhi syarat kriteria yang ditentukan.

Tabel 2. Validitas butir soal kelas XI

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PT-MEASURE CORR.	EXP.	EXACT MATCH OBS%	EXP%	Item
14	30	164	90.94	2.27	1.73	4.6	2.28	3.9	A-.14	.43	70.1	83.8	14
7	156	164	36.96	3.71	.99	.1	1.60	1.0	B .16	.18	95.1	95.1	7
13	142	164	48.86	2.42	1.00	.1	1.38	1.1	C .24	.29	86.6	86.6	13
27	163	164	15.29	10.08	1.02	.4	1.28	.6	D .02	.07	99.4	99.4	27
21	153	164	40.52	3.22	1.05	.3	1.26	.6	E .15	.21	93.3	93.3	21
12	101	164	65.34	1.80	1.14	1.9	1.04	.4	F .36	.44	64.6	70.5	12
29	115	164	60.63	1.88	1.13	1.7	1.03	.3	G .32	.41	75.0	73.5	29
4	114	164	60.98	1.87	1.12	1.6	1.00	.1	H .33	.41	69.5	73.1	4
16	155	164	38.26	3.52	.97	.0	1.11	.4	I .20	.19	94.5	94.5	16
11	136	164	52.07	2.22	1.09	.7	1.05	.3	J .25	.32	82.9	83.1	11
3	147	164	45.64	2.68	1.04	.3	1.03	.2	K .22	.26	89.6	89.6	3
19	158	164	33.83	4.24	1.03	.2	.89	.0	L .13	.16	96.3	96.3	19
17	154	164	39.44	3.36	1.01	.1	1.02	.2	M .19	.20	93.9	93.9	17
10	142	164	48.86	2.42	1.01	.1	.97	.0	N .27	.29	86.6	86.6	10
20	158	164	33.83	4.24	.99	.1	.99	.2	O .17	.16	96.3	96.3	20
2	140	164	49.99	2.34	.99	.0	.87	-.3	o .32	.30	86.0	85.4	2
22	84	164	70.75	1.78	.97	-.4	.98	-.2	n .49	.47	72.6	70.7	22
25	135	164	52.55	2.19	.98	-.2	.77	-.8	m .37	.33	82.3	82.5	25
18	155	164	38.26	3.52	.96	.0	.82	-.1	l .23	.19	94.5	94.5	18
23	148	164	44.90	2.74	.95	-.2	.63	-.9	k .32	.25	90.2	90.3	23
30	138	164	51.06	2.27	.95	-.4	.75	-.8	j .37	.31	84.8	84.2	30
5	98	164	66.31	1.79	.95	-.7	.90	-.8	i .49	.45	77.4	70.3	5
1	135	164	52.55	2.19	.95	-.4	.72	-1.0	h .39	.33	82.3	82.5	1
6	154	164	39.44	3.36	.93	-.2	.51	-.9	g .29	.20	93.9	93.9	6
8	139	164	50.53	2.31	.93	-.5	.82	-.5	f .36	.31	85.4	84.8	8
26	117	164	59.91	1.90	.90	-1.2	.78	-1.2	e .48	.40	76.2	74.3	26
9	160	164	29.53	5.13	.89	-.1	.25	-1.3	d .27	.13	97.6	97.6	9
15	144	164	47.65	2.51	.88	-.7	.51	-1.5	c .41	.28	87.8	87.8	15
28	69	164	75.58	1.82	.81	-2.5	.83	-1.6	b .62	.49	76.2	72.4	28
24	118	164	59.55	1.91	.76	-3.0	.60	-2.4	a .58	.40	80.5	74.6	24
MEAN	131.9	164.0	50.00	2.92	1.00	.0	.96	-.2			85.4	85.4	
S.D.	30.0	.0	14.90	1.58	.16	1.3	.36	1.1			9.3	9.1	

Berdasarkan tabel terlihat bahwa tidak semua item memenuhi ketiga kategori yang disyaratkan agar *item* tersebut valid, berikut uraiannya:

- I. Pada item (soal) nomor 7 dan nomor 9 hanya memenuhi kriteria ZSTD.
 - m. Pada item (soal) nomor 24 hanya memenuhi kriteria MNSQ dan Pt Mean Cor.
 - n. Pada item (soal) nomor 13, 27, 21, 12, 29, 4, 16, 11, 3, 19, 17, 10, 20, 2, 22, 25, 18, 23, 30, 1, 6, dan 8 hanya memenuhi kriteria MNSQ dan ZSTD.
 - o. Pada item (soal) nomor 5, 26, 15, dan 28 memenuhi kriteria MNSQ, ZSTD, dan Pt Mean Cor.

Meskipun hanya 4 item yang memenuhi ketiga kategori yang disyaratkan, *item* lainnya yang hanya memenuhi satu atau dua kategori tetap dikatakan valid. Sehingga dari 30 soal terdapat 29 soal yang valid dan satu butir soal yang tidak valid yaitu nomor 14, hal ini dikarenakan nilai dari MNSQ, ZSTD, dan Pt Mean Cor tidak ada yang memenuhi kriteria.

Reliabilitas

Reliabilitas adalah sama dengan konsistensi atau keajegan. Suatu instrumen evaluasi dikatakan mempunyai nilai reliabilitas tinggi, apabila tes yang dibuat mempunyai hasil yang konsisten dalam mengukur yang hendak diukur (Sukardi, 2011: 29).

Analisa reliabilitas soal dengan menggunakan aplikasi *Winsteps* dengan model analisis *Rasch*. Pada soal pilihan ganda dengan pendekatan sistemik kelas XI, data soal nomor 14 dihapus karena soal tersebut tidak valid. Sehingga data yang digunakan untuk menentukan nilai reliabilitas soal hanya berdasarkan 29 soal yang valid.

Tabel 3. Reliabilitas soal kimia Kelas XI

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INF IT		OUTF IT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	24.0	29.0	74.34	7.44				
S.D.	3.5	.0	14.23	4.36				
MAX.	29.0	29.0	103.64	18.59				
MIN.	15.0	29.0	50.91	4.29	.63	-1.5	.15	-1.0

REAL RMSE	8.81	TRUE SD	11.17	SEPARATION	1.27	Person RELIABILITY .62		
MODEL RMSE	8.62	TRUE SD	11.32	SEPARATION	1.31	Person RELIABILITY .63		
S.E. OF Person MEAN = 1.11								

Person RAW SCORE-TO-MEASURE CORRELATION = .94								
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .73								
SUMMARY OF 29 MEASURED Item								

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INF IT		OUTF IT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	135.4	164.0	50.00	2.96	.99	.0	1.07	.2
S.D.	23.7	.0	13.24	1.59	.09	1.0	.39	1.1
MAX.	163.0	164.0	78.11	10.08	1.18	2.6	2.31	2.6
MIN.	69.0	164.0	16.54	1.84	.78	-3.0	.29	-2.6

REAL RMSE	3.39	TRUE SD	12.80	SEPARATION	3.77	Item RELIABILITY .93		
MODEL RMSE	3.36	TRUE SD	12.81	SEPARATION	3.81	Item RELIABILITY .94		
S.E. OF Item MEAN = 2.50								

Tabel 3 menunjukkan tabel ringkasan mengenai responden (siswa) pada bagian atas dan tabel ringkasan mengenai *item* (soal) pada bagian bawah. Berdasarkan tabel di tersebut, untuk soal kelas XI reliabilitas responden dapat diketahui dari nilai *person reliability* sebesar 0,62 atau 62% yang menunjukkan bahwa reliabilitas responden lemah dengan kata lain dari 164 peserta didik yang mengerjakan soal kesalahan menjawab setiap peserta didik pada soal yang berbeda meskipun jumlah skor mereka sama. Reliabilitas butir soal dapat diketahui dari nilai *item reliability* sebesar 0,93 atau 93% yang menunjukkan bahwa reliabilitas butir soal bagus sekali dengan kata lain soal dapat mengukur kemampuan siswa dengan baik.

Menurut Oemar Hamalik (1989: 37) tes yang baik jika memenuhi persyaratan validitas, reliabilitas, dan objektifitas. Berdasarkan uji validitas, reliabilitas dan tingkat kesukaran bahwa soal evaluasi berbasis pendekatan sistemik yang telah dikembangkan dapat digunakan sebagai alat evaluasi dan merupakan alat evaluasi yang baik.

Soal evaluasi berbasis pendekatan sistemik yang dihasilkan dari penelitian ini memiliki kelebihan antara lain :

1. Materi yang digunakan sesuai dengan kurikulum 2013.
2. Soal evaluasi ini dapat membantu peserta didik berfikir konstruktif yaitu mengingat kembali materi yang telah dipelajarinya.
3. Soal evaluasi ini dapat digunakan oleh guru sebagai ulangan harian.
4. Soal evaluasi ini telah diujikan kepada peserta didik kelas XI dan hasil yang diperoleh sangat baik.

SIMPULAN

1. Telah berhasil dikembangkan suatu alat evaluasi pembelajaran kimia berupa soal pilihan ganda sistemik yang berkualitas sangat baik berdasarkan penilaian reviewer.
2. Hasil ujicoba soal menunjukkan bahwa soal yang dikembangkan valid dan mempunyai reliabilitas bagus sekali.

DAFTAR PUSTAKA

- Bambang Sumintono dan Wahyu Widhiarso. (2014). *Aplikasi Model Rasch*. Jakarta: Trim Komunikasi Publishing House
- Chiheb, Raddouan., Faizi, Rdouan., Afia, Abdellatif El. (2011). Using Objective Online Testing Tools To Assess Students' Learning: Potential and Limitation. *Journal of Theoretical and Applied Information Technology* Ed. 8, Vol. 24 (1), p. 69-72
- Djemari Mardapi. (2008). *Teknik Penyusunan Instrumen Tes dan Nontes*. Jogjakarta: Mitra Cendikia
- Fahmy, A.F.M. dan Lagowski, J.J. (2006). Systematic Multiple Choice Question in Chemistry. *19th ICCE Proceeding, Seoul, Aug. 12-17 2006*
- Gall, J., Borg, W. & Gall, M. (2003). *Educational Research: An Introduction (7th ed)*. Boston:Pearson Education
- Lagowski, J.J. (2008). SATL, Learning Theory, and the Physiology of Learning. *20th ICCE Proceeding, Mauritius, Aug.3-8 2008*.
- Oemar Hamalik. (1989). *Teknik Pengukuran dan Evaluasi Pendidikan*. Bandung: Mandar Maju
- Sukardi. (2011). *Evaluasi Pendidikan Prinsip & Operasional*. Jakarta : Bumi Aksara
- Wina Sanjaya. (2008). *Kurikulum dan Pembelajaran*. Bandung: Kencana Prenada Media Group.
- Zimmaro, Dawn M. (2004). *Writing Good Multiple Choice-Exam*. Austin: DIIA The University of Texas at Austin.